

---

# Benchmark on Anisotropic Problems

## Nonlinear monotone finite volume method

Daniil Svyatskiy

*Los Alamos National Laboratory, MS B284, Los Alamos NM, 87545  
dasyyat@lanl.gov*

---

*ABSTRACT.* We consider a nonlinear finite volume (FV) method for stationary diffusion equation. We prove that the method is monotone, i.e. it preserves positivity of analytical solutions on arbitrary triangular meshes for strongly anisotropic and heterogeneous full tensor coefficients. The method is extended to regular star-shaped polygonal meshes and isotropic heterogeneous coefficients. The method has been developed in collaboration with K. Lipnikov, M. Shashkov and Y. Vassilevski and is based on ideas proposed by C. Le Potier.

*KEYWORDS:* Anisotropy benchmark, finite volume, nonlinear iterations.

---

### 1. Presentation of the scheme

Let  $\Omega$  be a two-dimensional polygonal domain  $\Omega$  with boundary  $\Gamma = \Gamma_N \cup \Gamma_D$ . We consider a diffusion problem for unknown  $u$ :

$$\begin{aligned} -\operatorname{div} \mathbb{K} \nabla u &= f && \text{in } \Omega \\ u &= g_D && \text{on } \Gamma_D \\ -\mathbb{K} \frac{\partial u}{\partial \mathbf{n}} &= g_N && \text{on } \Gamma_N \end{aligned} \tag{1}$$

where  $\mathbb{K} = \mathbb{K}^T > 0$  is the diffusion tensor,  $f$  is the source function, and  $\mathbf{n}$  is the exterior normal vector.

We assume that the tensor  $\mathbb{K}$  is constant inside each cell. Let  $\mathbf{q} = -\mathbb{K} \nabla u$  denote the diffusion flux which satisfies the mass balance equation:

$$\operatorname{div} \mathbf{q} = f \quad \text{in } \Omega. \tag{2}$$

Let  $\mathcal{T}$  be a conformal partition composed of  $N_T$  cells  $T$ . Integrating mass balance equation [2] over a cell  $T$  and using the Green formula we get:

$$\int_{\partial T} \mathbf{q} \cdot \mathbf{n} \, ds = \int_T f \, dx, \quad \forall T \in \mathcal{T}, \tag{3}$$

where  $\mathbf{n}$  denotes the outer unit normal to  $\partial T$ . Let  $e$  denote an edge of triangle  $T$  and  $\mathbf{n}_e$  be the corresponding outward normal vector. Hereafter, it will be convenient to assume that  $|\mathbf{n}_e| = |e|$  where  $|e|$  denotes the length of edge  $e$ . Equation [3] becomes

$$\sum_{e \in \partial T} \mathbf{q}_e \cdot \mathbf{n}_e = \int_T f \, dx, \quad \forall T \in \mathcal{T}, \quad [4]$$

where  $\mathbf{q}_e$  is the average flux density for edge  $e$ :

$$\mathbf{q}_e = \frac{1}{|e|} \int_e \mathbf{q} \, ds.$$

The FV schemes differ by approximations for the fluxes  $\mathbf{q}_e$ . In this method we use a two-point flux approximation. For each cell  $T$ , we assign one degree of freedom  $U_T$  for unknown function  $u$ . Let  $U$  be the vector of discrete unknowns. The two-point flux approximation uses only two degrees of freedom  $U_{T+}$  and  $U_{T-}$  corresponding to cells  $T_+$  and  $T_-$  that share the edge  $e$ . For simplicity we shall write  $U_+$  instead of  $U_{T+}$ . The general form for the two-point flux is as follows:

$$\mathbf{q}_e^h \cdot \mathbf{n}_e = A_e^+ U_+ - A_e^- U_-,$$

where  $A_e^+$  and  $A_e^-$  are coefficients which will be defined in the next section. Substituting discrete approximation  $\mathbf{q}_e^h$  for  $\mathbf{q}_e$  in [4], we obtain a system of  $N_T$  equations with  $N_T$  unknowns  $U_T$ .

### 1.1. Triangular meshes

We consider a nonlinear two-point flux approximation where coefficients  $A_e^+$  and  $A_e^-$  depend on concentration. We begin with the physical meaning of discrete unknowns. The discrete concentration  $U_T$  approximates the unknown continuous function  $u$  at a point  $\mathbf{x}_T$  inside triangle  $T$ . We shall refer to this point as the *collocation point*. Denoting the vertices of this triangle by  $\mathbf{v}_1$ ,  $\mathbf{v}_2$  and  $\mathbf{v}_3$ , we define the collocation point as follows:

$$\mathbf{x}_T = \sum_{i=1}^3 \mathbf{v}_i \lambda_i, \quad \lambda_i = \frac{|\mathbf{n}_{\alpha(i)}|_{\mathbb{K}}}{\sum_{j=1}^3 |\mathbf{n}_{\alpha(j)}|_{\mathbb{K}}}, \quad [5]$$

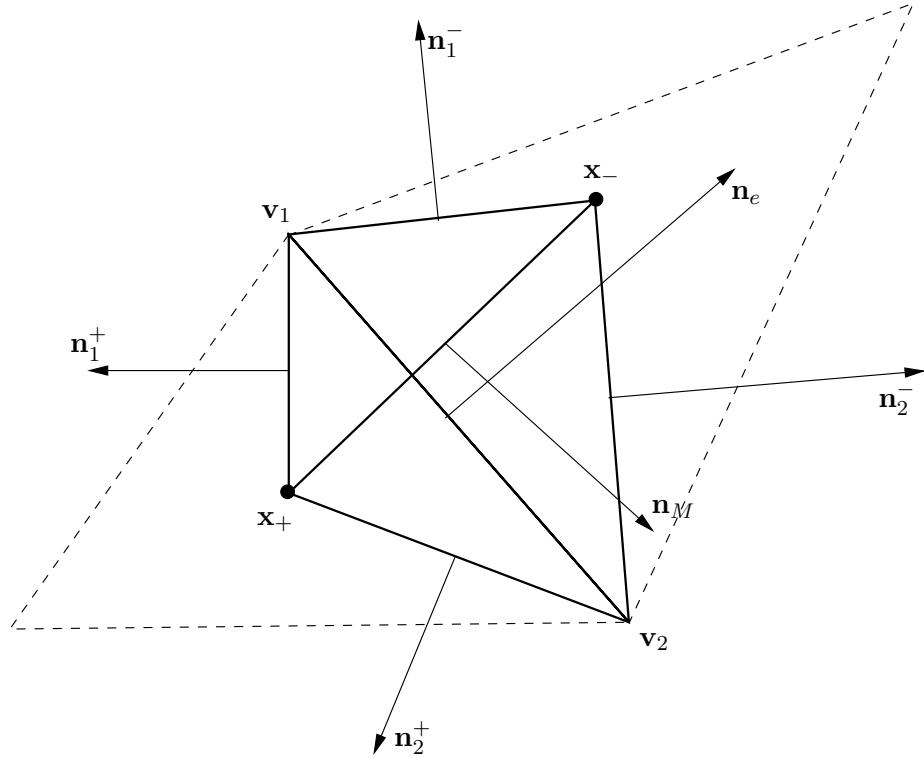
where  $|\mathbf{n}|_{\mathbb{K}} = (\mathbb{K} \mathbf{n} \cdot \mathbf{n})^{1/2}$  is the length of vector  $\mathbf{n}$  in metric  $\mathbb{K}$  induced by the diffusion tensor in triangle  $T$  and  $\alpha(i)$  denotes the edge opposite to vertex  $\mathbf{v}_i$ . The reason for such a choice of coordinates  $\lambda_i$  will be explained later.

Let us consider an interior mesh edge  $e$  with end points  $\mathbf{v}_1$  and  $\mathbf{v}_2$  shared by two triangles  $T_+$  and  $T_-$ . Let  $\mathbb{K}_+$  and  $\mathbb{K}_-$  be the values of diffusion tensor in triangles  $T_+$  and  $T_-$ , respectively. For simplicity we assume that  $\mathbb{K}_+ = \mathbb{K}_- = \mathbb{K}$ . The case  $\mathbb{K}_+ \neq \mathbb{K}_-$  is considered in [LIP 07]. Similarly, we denote the collocation points for

these triangles by  $\mathbf{x}_+$  and  $\mathbf{x}_-$  (see Figure 1). We assume that the normal vector  $\mathbf{n}_e$  is outward for triangle  $T_+$ .

Let  $T_i$ ,  $i = 1, 2$ , be the triangle with vertices  $\mathbf{x}_+$ ,  $\mathbf{x}_-$ , and  $\mathbf{v}_i$ . For triangle  $T_1$ , we denote the normal vectors to its edges by  $\mathbf{n}_1^+$ ,  $\mathbf{n}_1^-$  and  $\mathbf{n}_M$  as shown in Figure 1. We assume again that the length of these vectors equals the length of the corresponding edge, i.e.  $|\mathbf{n}_1^\pm| = |\mathbf{v}_1 - \mathbf{x}_\pm|$  and  $|\mathbf{n}_M| = |\mathbf{x}_+ - \mathbf{x}_-|$ . In a similar way we define normals  $\mathbf{n}_2^\pm$  to edges of triangle  $T_2$ . The following identities hold:

$$\mathbf{n}_1^+ + \mathbf{n}_1^- + \mathbf{n}_M = 0 \quad \text{and} \quad \mathbf{n}_2^+ + \mathbf{n}_2^- - \mathbf{n}_M = 0. \quad [6]$$



**Figure 1.** Interior edge  $e$  with end points  $\mathbf{v}_1$  and  $\mathbf{v}_2$ . The collocation points  $\mathbf{x}_+$  and  $\mathbf{x}_-$  are marked by solid balls. The triangles  $T_+$  and  $T_-$  are marked by dashed lines.

The Green formula for triangle  $T_1$  and definition of flux  $\mathbf{q}$  yield:

$$\int_{T_1} \mathbb{K}^{-1} \mathbf{q} dx = - \int_{\partial T_1} u \mathbf{n} ds. \quad [7]$$

Applying the mid-point (second-order) quadrature rule for both integrals, we obtain

$$-|T_1|\mathbb{K}^{-1}\mathbf{q}_{e,1}^h = \frac{U_1 + U_+}{2} \mathbf{n}_1^+ + \frac{U_1 + U_-}{2} \mathbf{n}_1^- + \frac{U_+ + U_-}{2} \mathbf{n}_M$$

where  $U_1$ ,  $U_+$  and  $U_-$  are the values of concentration  $c$  at points  $\mathbf{v}_1$ ,  $\mathbf{x}_+$ , and  $\mathbf{x}_-$ , respectively. Only concentrations  $U_{\pm}$  are our discrete unknowns. The concentration  $U_1$  will be eliminated later. Using identity [6], we get

$$\mathbf{q}_{e,1}^h = \frac{1}{2|T_1|}\mathbb{K}(U_+\mathbf{n}_1^- + U_-\mathbf{n}_1^+ - U_1(\mathbf{n}_1^+ + \mathbf{n}_1^-)). \quad [8]$$

Now we apply the same derivations to triangle  $T_2$  to obtain the second formula for the flux density:

$$\mathbf{q}_{e,2}^h = \frac{1}{2|T_2|}\mathbb{K}(U_+\mathbf{n}_2^- + U_-\mathbf{n}_2^+ - U_2(\mathbf{n}_2^+ + \mathbf{n}_2^-)). \quad [9]$$

Given two flux density approximations [8] and [9], we seek for the discrete flux  $\mathbf{q}_e^h \cdot \mathbf{n}_e$  through edge  $e$  as their linear combination:

$$\mathbf{q}_e^h \cdot \mathbf{n}_e = \mu_1 \mathbf{q}_{e,1}^h \cdot \mathbf{n}_e + \mu_2 \mathbf{q}_{e,2}^h \cdot \mathbf{n}_e, \quad [10]$$

where  $\mu_1$  and  $\mu_2$  are positive unknown coefficients. The approximation of flux density yields

$$\mu_1 + \mu_2 = 1. \quad [11]$$

Taking  $\mu_1$  and  $\mu_2$  as follows:

$$\mu_1 = \frac{U_2/|T_2|}{U_1/|T_1| + U_2/|T_2|} \quad \text{and} \quad \mu_2 = \frac{U_1/|T_1|}{U_1/|T_1| + U_2/|T_2|}. \quad [12]$$

we obtain the definition of the discrete flux written in the following form:

$$\mathbf{q}_e^h \cdot \mathbf{n}_e = A_e^+(U)U_+ - A_e^-(U)U_-, \quad [13]$$

where

$$\begin{aligned} A_e^+(U) &= \frac{\mu_1}{2|T_1|} \mathbf{n}_1^- \cdot \mathbb{K} \mathbf{n}_e + \frac{\mu_2}{2|T_2|} \mathbf{n}_2^- \cdot \mathbb{K} \mathbf{n}_e \\ A_e^-(U) &= -\frac{\mu_1}{2|T_1|} \mathbf{n}_1^+ \cdot \mathbb{K} \mathbf{n}_e - \frac{\mu_2}{2|T_2|} \mathbf{n}_2^+ \cdot \mathbb{K} \mathbf{n}_e. \end{aligned} \quad [14]$$

The coefficients  $A_e^+$  and  $A_e^-$  depend on concentrations  $U_1, U_2$ , i.e. flux [13] is *non-linear*. The unknown concentrations  $U_1$  and  $U_2$  must be approximated using the original degrees of freedom, i.e. concentrations at collocation points. The total number of

collocation points is  $N_T$  which leave enough flexibility for accurate approximation of these concentrations. We consider two interpolation methods.

The first interpolation method uses three collocation points closest to  $\mathbf{v}_1$  that form a imaginary non-degenerate triangle  $\tilde{T}$  containing  $\mathbf{v}_1$ . We denote these points by  $\mathbf{x}_{T_j}$ ,  $j = 1, 2, 3$ . The linear interpolation over this triangle gives a second-order approximation for  $U_1$  for smooth solutions [LEP 05]:

$$U_1 = \sum_{j=1}^3 U(\mathbf{x}_{T_j}) \tilde{\lambda}_j \quad [15]$$

where  $\tilde{\lambda}_j$ ,  $j = 1, 2, 3$ , are the barycentric coordinates of point  $\mathbf{v}_1$  in triangle  $\tilde{T}$ . Note that  $0 \leq \tilde{\lambda}_j \leq 1$ . We found out that this interpolation method is not robust for problems with strong anisotropy and/or solutions with sharp gradients.

Second interpolation method uses the inverse distance weighting [SHE 68] of values  $U(\mathbf{x}_T)$  for all triangles  $T \in \mathcal{T}$  that have  $\mathbf{v}_1$  as a vertex. Let  $\mathcal{U}(\mathbf{v}_1)$  be the collection of these triangles. Then

$$U_1 = \sum_{T \in \mathcal{U}(\mathbf{v}_1)} U(\mathbf{x}_T) w_T, \quad w_T = \frac{|\mathbf{x}_T - \mathbf{v}_1|^{-1}}{\sum_{T' \in \mathcal{U}(\mathbf{v}_1)} |\mathbf{x}_{T'} - \mathbf{v}_1|^{-1}}. \quad [16]$$

Note that  $0 \leq w_T \leq 1$ . We shall use this fact later. The same interpolation methods are used for approximating  $U_2$ .

## 1.2. Discrete system and its iterative solution

The substitution of flux [14] into equation [4] yields the nonlinear system:

$$\mathbb{A}(C)C = F. \quad [17]$$

To solve system [17] we use the Picard iterative method: choose a small value  $\varepsilon_{non} > 0$  and initial vector  $U^0 \in \mathfrak{R}^{N_T}$  with positive entries,  $U_i^0 \geq 0$ ,  $i = 1, \dots, N_T$ , and repeat for  $k = 1, 2, \dots$ ,

- 1) solve  $\mathbb{A}(U^{k-1})U^k = F$ ,
- 2) stop if  $\|\mathbb{A}(U^k)U^k - F\| \leq \varepsilon_{non} \|\mathbb{A}(U^0)U^0 - F\|$ .

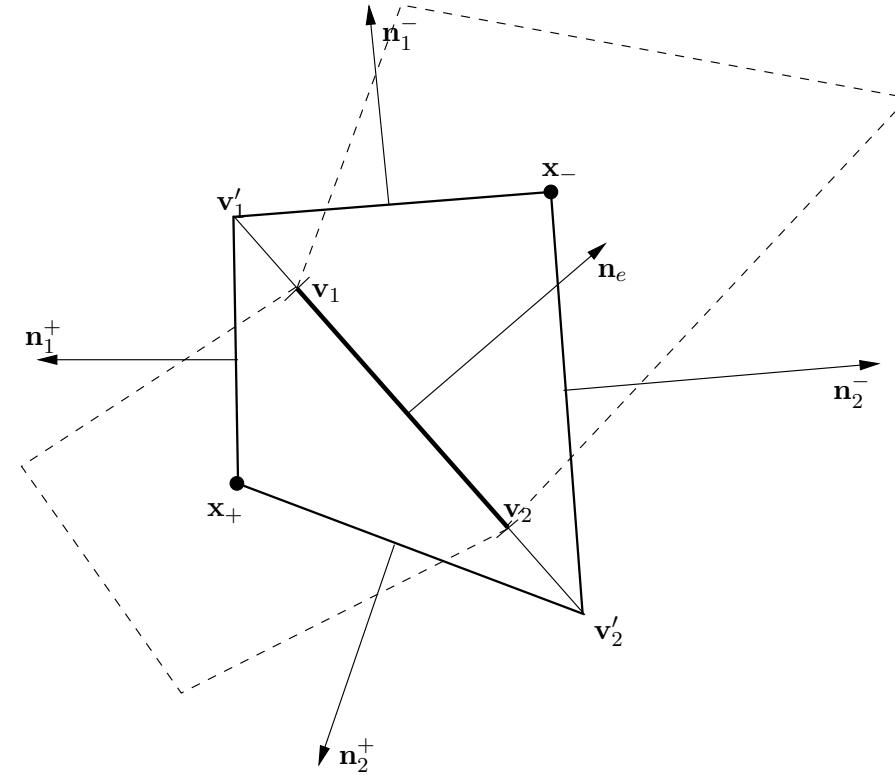
The linear system with non-symmetric matrix  $\mathbb{A}(U^{k-1})$  is solved by the Bi-Conjugate Gradient Stabilized (BCGStab) method. The BCGStab iterations are terminated when the relative norm of the initial residual becomes smaller than  $\varepsilon_{lin}$ . Matrix  $\mathbb{A}(U^{k-1})$  has at most four non-zeros in each row.

It has been proved in [LIP 07] that if the collocation points defined by [5] then the matrix  $\mathbb{A}(U^{k-1})$  is monotone for any non-negative vector  $U^{k-1}$ . The proof is based

on the fact that the choice of  $\mathbf{x}_{T_i}$  yields positive  $A_e^+$  and  $A_e^-$  in [14] for any piecewise constant diffusion tensor. Then the solution  $U^k$  of  $\mathbb{A}(U^{k-1})U^k = F$  is a non-negative vector and the next matrix  $\mathbb{A}(U^k)$  is again monotone. Therefore,  $U_{T_i}^k \geq 0$  for all  $i$  and  $k$ .

## 2. Monotone nonlinear FV scheme on polygonal meshes

Construction of a nonlinear FV scheme on a polygonal mesh is similar to that on a triangular mesh. The main difficulty is to determine a position of collocation point inside each mesh cell such that the resulting system is monotone. For the triangular case it is proved that such points exist for any diffusion tensor and any geometry. For general polygonal meshes such points exist only for a restricted class of meshes and diffusion tensors. We modify the scheme to relax some of the restrictions.



**Figure 2.** Interval  $[v'_1; v'_2]$  containing the interior mesh edge  $e$  with end points  $v_1$  and  $v_2$ . The collocation points  $x_+$  and  $x_-$  are marked by solid balls. The quadrilaterals  $Q_+$  and  $Q_-$  are marked by dashed lines.

Let  $\mathbb{K}$  be an isotropic or slightly anisotropic heterogeneous diffusion tensor and  $\mathcal{Q}$  be a conformal polygonal mesh composed of  $N_{\mathcal{Q}}$  cells. We assume that the mesh is composed of *shape-regular and star-shaped* cells in the following sense:

- 1) For each polygonal cell  $Q \in \mathcal{Q}$ , we have

$$\frac{d(Q)}{\rho(Q)} \leq R_*,$$

where  $d(Q)$  is the diameter of  $Q$ ,  $\rho(Q)$  is the radius of maximal inscribed circle, and  $R_*$  is a constant independent of the mesh.

2) Each cell  $Q$  is star-shaped with respect to an interior point  $\mathbf{x}_Q$ , i.e. any ray emanating from this point intersects the boundary  $\partial Q$  at exactly one point. If geometry allows, we shall always place  $\mathbf{x}_Q$  at the center of mass of  $Q$ .

Given a two-point flux formula [13] we may follow the path described in the previous section to get a nonlinear system [17]. In order to guarantee positivity of coefficients in formula [13], we propose the following method. For an edge  $e \in \mathcal{E}_I$  with end points  $\mathbf{v}_1$  and  $\mathbf{v}_2$ , we define a minimal interval  $e' = [\mathbf{v}'_1; \mathbf{v}'_2]$  containing  $e$  such that

$$\mathbb{K}\mathbf{n}_i^- \cdot \mathbf{n}_e \geq 0 \quad \text{and} \quad \mathbb{K}\mathbf{n}_i^+ \cdot \mathbf{n}_e \leq 0, \quad i = 1, 2, \quad [18]$$

where  $\mathbf{n}_i^\pm$  are outward normals to edges of polygon  $\mathbf{v}'_1 \mathbf{x}_+ \mathbf{v}'_2 \mathbf{x}_-$  as shown in Figure 2. We may use formula [13] to calculate the flux density through  $e'$  and associate this flux density with the mesh edge  $e$ . The accuracy of such a modification depends on the ratio  $|e'|/|e|$  which is bounded for shape-regular polygonal meshes and isotropic heterogeneous tensors. The monotonicity of the matrix  $\mathbb{A}(U)$  for any non-negative vector  $U$  follows from [18].

The nonlinear FV method is monotone and conservative for arbitrary triangular meshes and arbitrary full tensor diffusion coefficients. It has the four-point stencil for triangular meshes and the five-point stencil for quadrilateral meshes. It gives the second-order convergence rate for the scalar unknown and the first-order convergence rate for the flux unknown in the case of smooth solutions. The price for these appealing features is the method non-linearity. The extension of this method for unstructured tetrahedral meshes is proposed in [KAP 07].

### 3. Numerical results

- **Test 1.1 Mild anisotropy,**  $u(x, y) = 16x(1-x)y(1-y)$ , min = 0, max = 1, regular triangular mesh, `mesh1`

i	nunkw	nnmat	sumflux	erl2	ergrad	ratio12	ratiograd
1	56	208	1.11e-15	4.05e-02	6.06e-01		
2	224	864	-8.52e-15	1.30e-02	3.03e-01	1.64e-00	1.00e-00
3	896	3520	-2.54e-14	3.73e-03	1.51e-01	1.80e-00	1.00e-00
4	3584	14208	-1.79e-14	1.01e-03	7.56e-02	1.88e-00	1.00e-00
5	14336	57088	-2.79e-13	2.66e-04	3.78e-02	1.92e-00	1.00e-00
6	57344	228864	-1.87e-12	6.80e-05	1.89e-02	1.96e-00	1.00e-00
7	229376	916480	-5.42e-12	1.72e-05	9.43e-03	1.98e-00	1.00e-00

`ocvl2`= 1.99e - 00, `ocvgradl2`= 1.00e - 00.

i	erflx0	erflx1	erfly0	erfly1	erflm	umin	umax
1	1.26e-04	1.26e-04	1.26e-04	1.26e-04	2.71e-01	1.05e-01	9.45e-01
2	1.26e-04	5.19e-05	1.87e-05	9.24e-05	1.52e-01	2.94e-02	9.88e-01
3	3.68e-06	2.33e-05	3.16e-05	1.20e-05	8.96e-02	7.61e-03	9.97e-01
4	7.74e-06	6.39e-06	5.84e-06	7.19e-06	4.98e-02	1.92e-03	9.99e-01
5	2.01e-06	4.08e-06	3.38e-06	1.32e-06	2.62e-02	4.82e-04	1.00e-00
6	2.68e-07	1.94e-07	2.38e-07	2.24e-07	1.34e-02	1.20e-04	1.00e-00
7	3.71e-07	2.63e-07	2.97e-07	4.05e-07	6.78e-03	3.01e-05	1.00e-00

- **Test 1.1 Mild anisotropy,**  $u(x, y) = 16x(1-x)y(1-y)$ , min = 0, max = 1, coarse(C) and fine (F) distorted quadrangular meshes , `mesh4_1` and `mesh4_2`

grid	nunkw	nnmat	sumflux	erl2	ergrad
C	289	1377	-1.81e-14	6.69e-02	
F	1089	5313	-9.68e-14	2.64e-02	

grid	erflx0	erflx1	erfly0	erfly1	erflm	umin	umax
C	1.53e-02	3.83e-02	5.18e-03	5.88e-02	7.33e-01	1.34e-02	1.11e+00
F	7.08e-03	1.27e-02	2.91e-03	1.69e-02	2.29e-01	3.61e-03	1.04e+00

- **Test 1.2 Mild anisotropy,**  $u(x, y) = \sin((1-x)(1-y)) + (1-x)^3(1-y)^2$ , min = 0, max = 1 + sin 1, regular triangular mesh, `mesh1`

i	nunkw	nnmat	sumflux	erl2	ergrad	ratio12	ratiograd
1	56	208	-1.60e-16	2.11e-02	2.53e-01		
2	224	864	-8.78e-15	5.35e-03	1.26e-01	1.97e-00	1.01e-00
3	896	3520	-7.46e-16	1.32e-03	6.24e-02	2.01e-00	1.01e-00
4	3584	14208	-3.78e-14	3.28e-04	3.11e-02	2.00e-00	1.00e-00
5	14336	57088	-3.38e-14	8.16e-05	1.55e-02	2.00e-00	1.00e-00
6	57344	228864	-4.41e-14	2.02e-05	7.75e-03	2.01e-00	1.00e-00
7	229376	916480	-2.13e-13	4.92e-06	3.87e-03	2.03e-00	1.00e-00

**ocvl2**=  $2.04e - 00$ , **ocvgradl2**=  $1.00e - 00$ .

i	erflx0	erflx1	erfly0	erfly1	erflm	umin	umax
1	2.22e-03	3.88e-02	3.71e-02	6.35e-02	2.16e-01	6.41e-03	1.39e+00
2	1.29e-03	9.63e-03	1.16e-02	2.00e-02	1.29e-01	1.55e-03	1.60e+00
3	5.25e-04	2.38e-03	3.48e-03	5.93e-03	6.90e-02	3.84e-04	1.72e+00
4	1.82e-04	5.89e-04	1.02e-03	1.73e-03	3.50e-02	9.58e-05	1.78e+00
5	5.86e-05	1.42e-04	2.89e-04	4.88e-04	1.78e-02	2.39e-05	1.81e+00
6	1.81e-05	3.36e-05	8.08e-05	1.36e-04	9.10e-03	5.98e-06	1.83e+00
7	5.76e-06	6.86e-06	2.19e-05	3.56e-05	4.48e-03	1.49e-06	1.83e+00

- **Test 1.2 Mild anisotropy**,  $u(x, y) = \sin((1-x)(1-y)) + (1-x)^3(1-y)^2$ , min = 0, max =  $1 + \sin 1$ , **locally refined non-conforming rectangular mesh**, mesh3

i	nunkw	nnmat	sumflux	erl2	ergrad	ratiol2	ratiograd
1	40	184	5.55e-16	9.61e-03			
2	160	768	-6.94e-17	2.80e-03		1.78e-00	
3	640	3136	-2.69e-15	8.07e-04		1.79e-00	
4	2560	12672	-9.61e-15	2.39e-04		1.75e-00	
5	10240	50944	-3.58e-14	6.26e-05		1.93e-00	

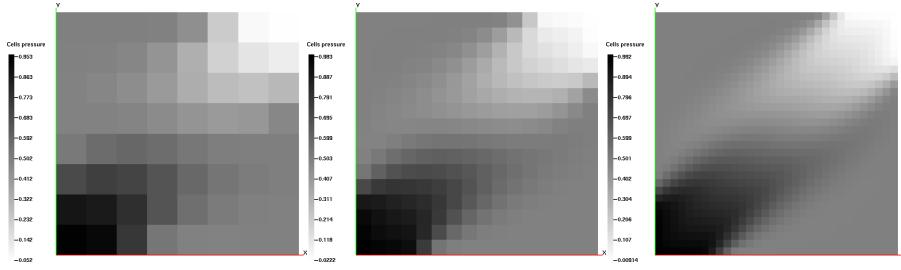
**ocvl2**=  $1.93e - 00$ .

i	erflx0	erflx1	erfly0	erfly1	erflm	umin	umax
1	9.11e-03	1.31e-02	1.47e-03	3.85e-02	2.07e-01	1.49e-02	1.66e+00
2	2.86e-03	1.75e-04	6.38e-04	1.21e-02	1.45e-01	3.71e-03	1.75e+00
3	7.14e-04	4.35e-04	1.39e-04	5.17e-03	7.41e-02	9.63e-04	1.79e+00
4	1.57e-04	1.60e-04	8.35e-05	1.26e-03	3.55e-02	2.43e-04	1.82e+00
5	3.56e-05	3.92e-05	2.50e-05	4.17e-04	2.02e-02	6.13e-05	1.83e+00

- **Test 3 Oblique flow**, min = 0, max = 1, **uniform rectangular mesh**, mesh2

Reference mesh is uniform rectangular mesh  $256 \times 256$ ,  $h = 3.90e-3$ .

i	nunkw	nnmat	sumflux	umin	umax
1	16	64	-1.53e-16	1.11e-01	8.88e-01
2	64	288	4.93e-16	5.20e-02	9.53e-01
3	256	1216	2.74e-16	2.22e-02	9.83e-01
4	1024	4992	-1.17e-15	9.14e-03	9.92e-01
5	4096	20224	4.15e-15	4.64e-03	9.96e-01
ref	65536	326656	-2.55e-14	1.28e-03	9.99e-01



**Figure 3.** Solutions for the oblique flow on  $\text{mesh2\_}i$  for  $i=2$  (left),  $i=3$  (center),  $i=4$  (right)

i	flux0	flux1	fluy0	fluy1	ener1	ener2	eren
1	-1.71e-01	1.72e-01	-1.66e-01	1.65e-01	2.33e-01	1.99e-01	1.45e-01
2	-1.89e-01	1.96e-01	-1.08e-01	1.01e-01	2.71e-01	2.69e-01	7.20e-03
3	-1.92e-01	2.00e-01	-9.88e-02	9.09e-02	3.07e-01	3.08e-01	3.56e-03
4	-1.94e-01	1.97e-01	-9.85e-02	9.56e-02	2.80e-01	3.01e-01	6.91e-02
5	-1.95e-01	1.96e-01	-9.73e-02	9.61e-02	2.62e-01	2.78e-01	5.69e-02
ref	-1.94e-01	1.94e-01	-9.79e-02	9.76e-02	2.45e-01	2.50e-01	1.94e-02

- **Test 4 Vertical fault**, min = 0., max = 1., **non-conforming rectangular mesh**,  $\text{mesh5}$

Reference mesh is uniform rectangular mesh  $320 \times 320$ ,  $h = 3.12e-3$ .

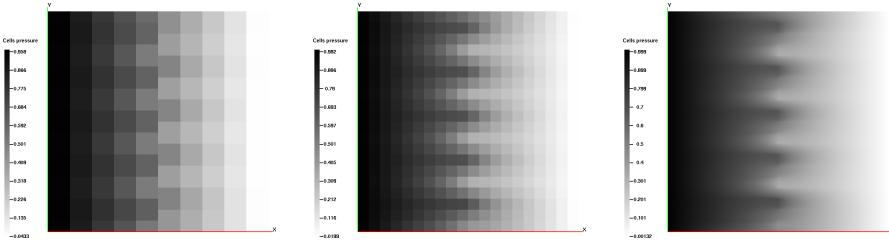
i	nunkw	nnmat	sumflux	umin	umax
1	105	503	2.53e-13	4.33e-02	9.58e-01
reg	400	1920	5.92e-13	1.99e-02	9.82e-01
ref	102400	510720	2.41e-11	1.32e-03	9.99e-01

i	flux0	flux1	fluy0	fluy1	ener1	ener2	eren
1	-4.32e+01	4.45e+01	-1.23e+00	2.32e-04		4.37e+01	
reg	-3.99e+01	4.26e+01	-2.68e+00	8.01e-04	3.97e+01	4.10e+01	3.22e-02
ref	-4.21e+01	4.44e+01	-2.33e+00	7.97e-04	4.32e+01	4.32e+01	5.92e-04

- **Test 6 Oblique drain**, min = -1.2, max = 0, **coarse oblique mesh**,  $\text{mesh6}$

grid	nunkw	nnmat	sumflux	erl2	ergrad
C	210	988	-2.91e-13	1.35e-04	
F	230	1128	-3.20e-14	1.57e-04	

grid	erflux0	erflux1	erfly0	erfly1	erflm	umin	umax
C	6.22e-04	8.13e-05	1.85e-03	1.30e-03	6.38e-01	-1.15e+00	-5.41e-02
F	6.18e-04	8.39e-05	1.86e-03	1.29e-03	6.38e-01	-1.15e+00	-5.41e-02



**Figure 4.** Solutions for the vertical fault on a non-conforming rectangular mesh (left), the square  $20 \times 20$  mesh (center), and the square  $320 \times 320$  mesh (right)

- **Test 7 Oblique barrier**, min =  $-5.575$ , max =  $0.575$ , **coarse oblique mesh** mesh6

h	nunkw	nnmat	sumflux	erl2	ergrad
3.29E-01	210	988	4.67e-14	4.98e-03	

erflx0	erflx1	erfly0	erfly1	erflm	umin	umax
1.75e+00	1.67e+00	2.36e-02	1.01e-02	7.09e+00	-5.54e+00	5.37e-01

- **Test 8 Perturbed parallelograms**, min =  $0$  , **perturbed parallelogram mesh** mesh8

nunkw	nnmat	sumflux	umin	umax
121	534	-2.22e-16	3.05e-15	9.42e-02

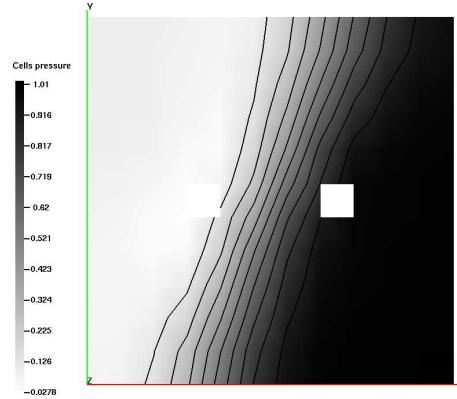
flux0	flux1	fluy0	fluy1
0.00e+00	0.00e+00	4.99e-01	5.01e-01

4.14e-15	1.83e-14	1.29e-11	7.76e-08	2.60e-05	9.12e-03	3.12e-05	1.46e-08	1.42e-10	5.80e-13	7.75e-14
1.28e-14	5.33e-14	3.69e-11	2.06e-07	9.02e-05	2.69e-02	8.98e-05	3.82e-08	4.10e-10	1.97e-12	2.17e-13
1.95e-14	9.36e-14	6.31e-11	2.86e-07	1.19e-04	4.48e-02	1.15e-04	6.42e-08	8.29e-10	3.29e-12	1.70e-13
2.91e-14	1.35e-13	9.30e-11	4.73e-07	1.25e-04	6.43e-02	1.25e-04	1.03e-07	1.06e-09	4.04e-12	1.66e-13
3.23e-14	1.22e-13	1.14e-10	4.53e-07	1.37e-04	7.89e-02	1.35e-04	1.22e-07	1.42e-09	6.61e-12	1.40e-13
3.62e-14	8.89e-14	1.02e-10	4.56e-07	1.30e-04	9.42e-02	1.04e-04	1.86e-07	1.65e-09	6.01e-12	1.32e-13
4.22e-14	6.88e-14	1.20e-10	3.34e-07	8.89e-05	8.19e-02	1.02e-04	2.44e-07	1.47e-09	5.80e-12	1.05e-13
2.71e-14	6.99e-14	1.26e-10	2.55e-07	8.50e-05	6.42e-02	1.03e-04	3.17e-07	1.04e-09	6.00e-12	9.07e-14
1.51e-14	8.62e-14	8.04e-11	2.22e-07	1.00e-05	4.68e-02	8.12e-05	2.73e-07	9.09e-10	6.13e-12	5.89e-14
9.55e-15	3.60e-14	4.56e-11	1.39e-07	1.09e-04	2.55e-02	4.89e-05	1.54e-07	5.53e-10	4.76e-12	3.37e-14
3.05e-15	1.08e-14	1.51e-11	5.46e-08	3.27e-05	9.15e-03	1.72e-05	5.99e-08	1.67e-10	1.33e-12	1.26e-14

Values of discrete solution obtained on mesh8 ( $11 \times 11$ ).

- **Test 9 Anisotropy with wells** , min =  $0$ , max =  $1$  , **square uniform mesh** mesh9

nunkw	nnmat	sumflux	umin	umax
119	539	1.80e-16	1.83e-02	1.01e+00



**Figure 5.** Solution for anisotropic problem with wells.

	7.01e-02	9.35e-02	1.93e-01	4.19e-01	7.21e-01	9.20e-01	1.01e+00	1.01e+00	1.01e+00	1.01e+00	1.01e+00
6.37e-02	7.07e-02	1.38e-01	3.03e-01	6.26e-01	8.63e-01	9.84e-01	1.01e+00	1.01e+00	1.01e+00	1.01e+00	1.01e+00
6.18e-02	5.97e-02	8.13e-02	2.01e-01	5.19e-01	7.98e-01	9.67e-01	1.00e+00	1.01e+00	1.01e+00	1.01e+00	1.01e+00
6.43e-02	5.26e-02	4.49e-02	1.03e-01	4.11e-01	7.25e-01	9.41e-01	9.92e-01	1.01e+00	1.01e+00	1.01e+00	1.01e+00
7.25e-02	5.44e-02	2.45e-02	2.48e-02	3.00e-01	6.14e-01	9.15e-01	9.87e-01	1.00e+00	1.01e+00	1.01e+00	1.01e+00
8.17e-02	6.81e-02	3.41e-02	1.24e-01	4.68e-01	8.33e-01	9.87e-01	1.00e+00	1.01e+00	1.01e+00	1.01e+00	1.01e+00
8.83e-02	8.36e-02	7.47e-02	1.84e-02	6.74e-02	3.29e-01	6.58e-01	9.71e-01	1.00e+00	1.00e+00	1.01e+00	
9.05e-02	8.75e-02	8.59e-02	4.12e-02	5.74e-02	2.50e-01	5.32e-01	8.54e-01	9.62e-01	9.98e-01	1.01e+00	
9.11e-02	8.93e-02	8.60e-02	5.91e-02	5.87e-02	1.89e-01	4.20e-01	7.32e-01	8.95e-01	9.80e-01	1.00e+00	
9.15e-02	9.05e-02	8.81e-02	6.76e-02	6.26e-02	1.43e-01	3.39e-01	6.07e-01	8.28e-01	9.42e-01	9.86e-01	
9.16e-02	9.10e-02	9.09e-02	7.39e-02	6.33e-02	1.09e-01	2.68e-01	5.06e-01	7.42e-01	8.95e-01	9.69e-01	

Values of discrete solution obtained on mesh9 (11 × 11).

#### 4. Comments on the results

##### Test 1.1, 1.2

- Triangular meshes (1.1, 1.2) – solution satisfies discrete maximum principle, the second order of convergence. The discrete gradient is calculated based on the prolongation of the discrete flux function using the lowest order Raviart-Thomas basis function. The first order of convergence for the discrete gradient.
- Distorted quadrangular meshes (1.1) – positivity discrete solution, small overshoots.
- Locally refined non-conforming rectangular mesh (1.2) – solution satisfies discrete maximum principle, second order of convergence.

**Test 2.** The main property of the method is to preserve positivity of the solution. The method is applicable to the problems that result in algebraic systems with positive right hand side and positive solutions, otherwise nonlinear iteration may diverge. It happens in this test.

**Test 3, 4.** The discrete solution satisfies the discrete maximum principle. In order to obtain ener1 the discrete gradient function should be calculated. The discrete gra-

dient function is not involved in the discretization so we need additional and artificial procedures to calculate it. There are two options. In the case of conformal quadrilateral mesh we can prolongate the flux function inside a cell using Raviart-Thomas basis functions. The main disadvantage of this approach consists the fact that in order to calculate the discrete gradient we need to multiply by  $\mathbb{K}^{-1}$ . If  $K$  is highly anisotropic it results in large errors for the discrete gradient. This effect is observed in Test 3. Following the second approach we can calculate the discrete gradient using degrees of freedom at collocation points and their interpolation at vertices, but the errors of interpolation procedures will significantly affect on the value of ener1. In Test 3 the results obtained by the second approach are presented. In Test4 we use Raviart-Thomas basis functions to calculate ener1.

**Test 5.** See comments for Test 2. Moreover, the generalization of the method for quadrilateral meshes works in case of isotropic or slightly anisotropic diffusion tensor. Otherwise an artificial edge extention can be large and it dramatically affects on the convergence. In this test edge extentions are about 15.

**Test 6.** Since solution is negative the method is applied to the problem with a solution  $\tilde{u}(x) = -u(x)$  and then multiplied by  $-1$ .

**Test 7.** Since solution is negative the method is applied to the problem with a solution  $\tilde{u}(x) = u(x) + 5.575$  and then subtract 5.575.

**Test 8.** The discrete solution is non-negative.

**Test 9.** The discrete solution is non-negative. Small overshoot is observed, see Figure 6.

## 5. References

- [KAP 07] KAPYRIN I., “A family of monotone methods for the numerical solution of three-dimensional diffusion problems on unstructured tetrahedral meshes”, *DOKLADY MATHEMATICS*, vol. 76, 2007, p. 734-738.
- [LEP 05] LEPOTIER C., “Schema volumes finis monotone pour des operateurs de diffusion fortement anisotropes sur des maillages de triangle non structures”, *C.R.Acad. Sci. Paris*, vol. Ser.I 341, 2005, p. 787–792.
- [LIP 07] LIPNIKOV K., SHASHKOV M., SVYATSKIY D., VASSILEVSKI Y., “Monotone finite volume schemes for diffusion equations on unstructured triangular and shape-regular polygonal meshes”, *Journal of Computational Physics*, vol. 227, 2007, p. 492–512.
- [SHE 68] SHEPARD D., “A two-dimensional interpolation function for irregularly spaced data”, *Proceedings of the 23d ACM National Conference*, 1968.